

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: HASTINGS, GREGG
- (ii) TITLE OF INVENTION: HUMAN NEURONAL ATTACHMENT FACTOR-1
- (iii) NUMBER OF SEQUENCES: 17
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: HUMAN GENOME SCIENCES, INC.
  - (B) STREET: 9410 KEY WEST AVENUE
  - (C) CITY: ROCKVILLE
  - (D) STATE: MD
  - (E) COUNTRY: USA
  - (F) ZIP: 20850
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: US
  - (B) FILING DATE: 11-FEB-1997
  - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: BROOKES, ANDERS A.
  - (B) REGISTRATION NUMBER: 36,373
  - (C) REFERENCE/DOCKET NUMBER: PF226
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: (301) 309-8504
  - (B) TELEFAX: (301) 309-8512

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1105 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 19..1011

## (ix) FEATURE:

(A) NAME/KEY: sig\_peptide  
(B) LOCATION: 19..87

## (ix) FEATURE:

(A) NAME/KEY: mat\_peptide  
(B) LOCATION: 88..1011

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CGCTGCTCCT GCCGGGTG ATG GAA AAC CCC AGC CCG GCC GCC GCC CTG GGC	51
Met Glu Asn Pro Ser Pro Ala Ala Ala Leu Gly	
-23 -20 -15	
AAG GCC CTC TGC GCT CTC CTC CTG GCC ACT CTC GGC GCC GCC GGC CAG	99
Lys Ala Leu Cys Ala Leu Leu Leu Ala Thr Leu Gly Ala Ala Gly Gln	
-10 -5 1	
CCT CTT GGG GGA GAG TCC ATC TGT TCC GCC AGA GCC CTG GCC AAA TAC	147
Pro Leu Gly Gly Glu Ser Ile Cys Ser Ala Arg Ala Leu Ala Lys Tyr	
5 10 15 20	
AGC ATC ACC TTC ACG GGC AAG TGG AGC CAG ACG GCC TTC CCC AAG CAG	195
Ser Ile Thr Phe Thr Gly Lys Trp Ser Gln Thr Ala Phe Pro Lys Gln	
25 30 35	
TAC CCC CTG TTC CGC CCC CCT GCC CAG TGG TCT TCG CTG CTG GGC GCC	243
Tyr Pro Leu Phe Arg Pro Pro Ala Gln Trp Ser Ser Leu Leu Gly Ala	
40 45 50	
GCG CAT AGC TCC GAC TAC AGC ATG TGG AGG AAG AAC CAG TAC GTC AGT	291
Ala His Ser Ser Asp Tyr Ser Met Trp Arg Lys Asn Gln Tyr Val Ser	
55 60 65	
AAC GGG CTG CGC GAC TTT GCG GAG CGC GGC GAG GCC TGG GCG CTG ATG	339
Asn Gly Leu Arg Asp Phe Ala Glu Arg Gly Glu Ala Trp Ala Leu Met	
70 75 80	
AAG GAG ATC GAG GCG GCG GGG GAG GCG CTG CAG AGC GTG CAC GCG GTG	387
Lys Glu Ile Glu Ala Ala Gly Glu Ala Leu Gln Ser Val His Ala Val	
85 90 95 100	
TTT TCG GCG CCC GCC GTC CCC AGC GGC ACC GGG CAG ACG TCG GCG GAG	435
Phe Ser Ala Pro Ala Val Pro Ser Gly Thr Gly Gln Thr Ser Ala Glu	
105 110 115	
CTG GAG GTG CAG CGC AGG CAC TCG CTG GTC TCG TTT GTG GTG CGC ATC	483
Leu Glu Val Gln Arg Arg His Ser Leu Val Ser Phe Val Val Arg Ile	
120 125 130	
GTG CCC AGC CCC GAC TGG TTC GTG GGC GTG GAC AGC CTG GAC CTG TGC	531
Val Pro Ser Pro Asp Trp Phe Val Gly Val Asp Ser Leu Asp Leu Cys	
135 140 145	

20250424

GAC GGG GAC CGT TGG CGG GAA CAG GCG GCG CTG GAC CTG TAC CCC TAC 579  
 Asp Gly Asp Arg Trp Arg Glu Gln Ala Ala Leu Asp Leu Tyr Pro Tyr  
 150 155 160

GAC GCC GGG ACG GAC AGC GGC TTC ACC TTC TCC TCC CCC AAC TTC GCC 627  
 Asp Ala Gly Thr Asp Ser Gly Phe Thr Phe Ser Ser Pro Asn Phe Ala  
 165 170 175 180

ACC ATC CCG CAG GAC ACG GTG ACC GAG ATA ACG TCC TCC TCT CCC AGC 675  
 Thr Ile Pro Gln Asp Thr Val Thr Glu Ile Thr Ser Ser Ser Pro Ser  
 185 190 195

CAC CCG GCC AAC TCC TTC TAC TAC CCG CGG CTG AAG GCC CTG CCT CCC 723  
 His Pro Ala Asn Ser Phe Tyr Tyr Pro Arg Leu Lys Ala Leu Pro Pro  
 200 205 210

ATC GCC AGG GTG ACA CTG GTG CGG CTG CGA CAG AGC CCC AGG GCC TTC 771  
 Ile Ala Arg Val Thr Leu Val Arg Leu Arg Gln Ser Pro Arg Ala Phe  
 215 220 225

ATC CCT CCC GCC CCA GTC CTG CCC AGC AGG GAC AAT GAG ATT GTA GAC 819  
 Ile Pro Pro Ala Pro Val Leu Pro Ser Arg Asp Asn Glu Ile Val Asp  
 230 235 240

AGC GCC TCA GTT CCA GAA ACG CCG CTG GAC TGC GAG GTC TCC CTG TGG 867  
 Ser Ala Ser Val Pro Glu Thr Pro Leu Asp Cys Glu Val Ser Leu Trp  
 245 250 255 260

TCG TCC TGG GGA CTG TGC GGA GGC CAC TGT GGG AGG CTC GGG ACC AAG 915  
 Ser Ser Trp Gly Leu Cys Gly Gly His Cys Gly Arg Leu Gly Thr Lys  
 265 270 275

AGC AGG ACT CGC TAC GTC CGG GTC CAG CCC GCC AAC AAC GGG AGC CCC 963  
 Ser Arg Thr Arg Tyr Val Arg Val Gln Pro Ala Asn Asn Gly Ser Pro  
 280 285 290

TGC CCC GAG CTC GAA GAA GAG GCT GAG TGC GTC CCT GAT AAC TGC GTC 1011  
 Cys Pro Glu Leu Glu Glu Glu Ala Glu Cys Val Pro Asp Asn Cys Val  
 295 300 305

TAAGACCAGA GCCCGCAGC CCCTGGGGCC CCCCGAGCC ATGGGGTGTC GGGGGCTCCT 1071  
 GTGCAGGCTC ATGCTGCAGG CGGCCGAGGG CACA 1105

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 331 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Glu Asn Pro Ser Pro Ala Ala Ala Leu Gly Lys Ala Leu Cys Ala  
 -23 -20 -15 -10  
 Leu Leu Leu Ala Thr Leu Gly Ala Ala Gly Gln Pro Leu Gly Gly Glu  
 -5 1 5  
 Ser Ile Cys Ser Ala Arg Ala Leu Ala Lys Tyr Ser Ile Thr Phe Thr  
 10 15 20 25  
 Gly Lys Trp Ser Gln Thr Ala Phe Pro Lys Gln Tyr Pro Leu Phe Arg  
 30 35 40  
 Pro Pro Ala Gln Trp Ser Ser Leu Leu Gly Ala Ala His Ser Ser Asp  
 45 50 55  
 Tyr Ser Met Trp Arg Lys Asn Gln Tyr Val Ser Asn Gly Leu Arg Asp  
 60 65 70  
 Phe Ala Glu Arg Gly Glu Ala Trp Ala Leu Met Lys Glu Ile Glu Ala  
 75 80 85  
 Ala Gly Glu Ala Leu Gln Ser Val His Ala Val Phe Ser Ala Pro Ala  
 90 95 100 105  
 Val Pro Ser Gly Thr Gly Gln Thr Ser Ala Glu Leu Glu Val Gln Arg  
 110 115 120  
 Arg His Ser Leu Val Ser Phe Val Val Arg Ile Val Pro Ser Pro Asp  
 125 130 135  
 Trp Phe Val Gly Val Asp Ser Leu Asp Leu Cys Asp Gly Asp Arg Trp  
 140 145 150  
 Arg Glu Gln Ala Ala Leu Asp Leu Tyr Pro Tyr Asp Ala Gly Thr Asp  
 155 160 165  
 Ser Gly Phe Thr Phe Ser Ser Pro Asn Phe Ala Thr Ile Pro Gln Asp  
 170 175 180 185  
 Thr Val Thr Glu Ile Thr Ser Ser Ser Pro Ser His Pro Ala Asn Ser  
 190 195 200  
 Phe Tyr Tyr Pro Arg Leu Lys Ala Leu Pro Pro Ile Ala Arg Val Thr  
 205 210 215  
 Leu Val Arg Leu Arg Gln Ser Pro Arg Ala Phe Ile Pro Pro Ala Pro  
 220 225 230  
 Val Leu Pro Ser Arg Asp Asn Glu Ile Val Asp Ser Ala Ser Val Pro  
 235 240 245  
 Glu Thr Pro Leu Asp Cys Glu Val Ser Leu Trp Ser Ser Trp Gly Leu  
 250 255 260 265

243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300

Cys Gly Gly His Cys Gly Arg Leu Gly Thr Lys Ser Arg Thr Arg Tyr  
 270 275 280

Val Arg Val Gln Pro Ala Asn Asn Gly Ser Pro Cys Pro Glu Leu Glu  
 285 290 295

Glu Glu Ala Glu Cys Val Pro Asp Asn Cys Val  
 300 305

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 36 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GCCATACGGG ATCCCCAGCC TCTTGGGGGA GAGTCC

36

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 35 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GGCATACGTC TAGATTAGAC GCAGTTATCA GGGAC

35

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 41 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GCCATACGGG ATCCGCCATC ATGGAAAACC CCAGCCCGGC C

41

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGCATACGTC TAGATTAGAC GCAGTTATCA GGGAC

35

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 392 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Pro	Thr	Gly	Thr	Gly	Cys	Val	Ile	Leu	Lys	Ala	Ser	Ile	Val	Gln	Lys
1				5					10					15	

Arg	Ile	Ile	Tyr	Phe	Gln	Asp	Glu	Gly	Ser	Leu	Thr	Lys	Lys	Leu	Cys
			20					25					30		

Glu	Gln	Asp	Pro	Thr	Leu	Asp	Gly	Val	Thr	Asp	Arg	Pro	Ile	Leu	Asp
		35					40					45			

Cys	Cys	Ala	Cys	Gly	Thr	Ala	Lys	Tyr	Arg	Leu	Thr	Phe	Tyr	Gly	Asn
	50					55					60				

Trp	Ser	Glu	Lys	Thr	His	Pro	Lys	Asp	Tyr	Pro	Arg	Arg	Ala	Asn	His
65					70				75					80	

Trp	Ser	Ala	Ile	Ile	Gly	Gly	Ser	His	Ser	Lys	Asn	Tyr	Val	Leu	Trp
				85					90					95	

Glu	Tyr	Gly	Gly	Tyr	Ala	Ser	Glu	Gly	Val	Lys	Gln	Val	Ala	Glu	Leu
			100					105						110	
Gly	Ser	Pro	Val	Lys	Met	Glu	Glu	Glu	Ile	Arg	Gln	Gln	Ser	Asp	Glu
		115					120					125			
Val	Leu	Thr	Val	Ile	Lys	Ala	Lys	Ala	Gln	Trp	Pro	Ser	Trp	Gln	Pro
	130					135					140				
Val	Asn	Val	Arg	Ala	Ala	Pro	Ser	Ala	Glu	Phe	Ser	Val	Asp	Arg	Thr
145					150					155					160
Arg	His	Leu	Met	Ser	Phe	Leu	Thr	Met	Met	Gly	Pro	Ser	Pro	Asp	Trp
				165					170					175	
Asn	Val	Gly	Leu	Ser	Ala	Glu	Asp	Leu	Cys	Thr	Lys	Glu	Cys	Gly	Trp
			180					185					190		
Val	Gln	Lys	Val	Val	Gln	Asp	Leu	Ile	Pro	Trp	Asp	Ala	Gly	Thr	Asp
		195					200					205			
Ser	Gly	Val	Thr	Tyr	Glu	Ser	Pro	Asn	Lys	Pro	Thr	Ile	Pro	Gln	Glu
	210					215					220				
Lys	Ile	Arg	Pro	Leu	Thr	Ser	Leu	Asp	His	Pro	Gln	Ser	Pro	Phe	Tyr
225					230					235					240
Asp	Pro	Glu	Gly	Gly	Ser	Ile	Thr	Gln	Val	Ala	Arg	Val	Val	Ile	Glu
				245					250					255	
Arg	Ile	Ala	Arg	Lys	Gly	Glu	Gln	Cys	Asn	Ile	Val	Pro	Asp	Asn	Val
		260						265					270		
Asp	Asp	Ile	Val	Ala	Asp	Leu	Ala	Pro	Glu	Glu	Lys	Asp	Glu	Asp	Asp
	275						280					285			
Thr	Pro	Glu	Thr	Cys	Ile	Tyr	Ser	Asn	Trp	Ser	Pro	Trp	Ser	Ala	Cys
	290					295					300				
Ser	Ser	Ser	Thr	Cys	Glu	Lys	Gly	Lys	Arg	Met	Arg	Gln	Arg	Met	Leu
305					310					315					320
Lys	Ala	Gln	Leu	Asp	Leu	Ser	Val	Pro	Cys	Pro	Asp	Thr	Gln	Asp	Phe
				325					330					335	
Gln	Pro	Cys	Met	Gly	Pro	Gly	Cys	Ser	Asp	Glu	Asp	Gly	Ser	Thr	Cys
			340					345					350		
Thr	Met	Ser	Glu	Trp	Ile	Thr	Trp	Ser	Pro	Cys	Ser	Val	Ser	Cys	Gly
		355					360					365			
Met	Gly	Met	Arg	Ser	Arg	Glu	Arg	Tyr	Val	Lys	Gln	Phe	Pro	Glu	Asp
	370					375					380				

Gly Ser Val Cys Met Leu Pro Thr  
385 390

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 52 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Cys	Ile	Tyr	Ser	Asn	Trp	Ser	Pro	Trp	Ser	Ala	Cys	Ser	Ser	Ser	Thr
1				5				10						15	
Cys	Glu	Lys	Gly	Lys	Arg	Met	Arg	Gln	Arg	Met	Leu	Lys	Ala	Gln	Leu
		20						25					30		
Asp	Leu	Ser	Val	Pro	Cys	Pro	Asp	Thr	Gln	Asp	Phe	Gln	Pro	Cys	Met
		35					40					45			
Gly	Pro	Gly	Cys												
		50													

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 53 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Cys	Thr	Met	Ser	Glu	Trp	Ile	Thr	Trp	Ser	Pro	Cys	Ser	Val	Ser	Cys
1				5				10						15	
Gly	Met	Gly	Met	Arg	Ser	Arg	Glu	Arg	Tyr	Val	Lys	Gln	Phe	Pro	Glu
		20						25					30		
Asp	Gly	Ser	Val	Cys	Met	Leu	Pro	Thr	Glu	Glu	Thr	Glu	Lys	Cys	Thr
		35					40					45			
Val	Asn	Glu	Glu	Cys											
		50													



## (2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 52 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Cys	Leu	Val	Thr	Glu	Trp	Gly	Glu	Trp	Asp	Asp	Cys	Ser	Ala	Thr	Cys
1				5				10						15	
Gly	Met	Gly	Met	Lys	Lys	Arg	His	Arg	Met	Val	Lys	Met	Ser	Pro	Ala
			20					25					30		
Asp	Gly	Ser	Met	Cys	Lys	Ala	Glu	Thr	Ser	Gln	Ala	Glu	Lys	Cys	Met
			35				40					45			
Met	Pro	Glu	Cys												
			50												

## (2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 51 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Cys	Leu	Leu	Ser	Pro	Trp	Ser	Glu	Trp	Ser	Asp	Cys	Ser	Val	Thr	Cys
1				5				10						15	
Gly	Lys	Gly	Met	Arg	Thr	Arg	Gln	Arg	Met	Leu	Lys	Ser	Leu	Ala	Glu
			20					25					30		
Leu	Gly	Asp	Cys	Asn	Glu	Asp	Leu	Glu	Gln	Ala	Glu	Lys	Cys	Met	Leu
			35				40					45			
Pro	Glu	Cys													
			50												

## (2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 52 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Cys	Glu	Leu	Ser	Glu	Trp	Ser	Gln	Trp	Ser	Glu	Cys	Asn	Lys	Ser	Cys
1				5					10					15	
Gly	Lys	Gly	His	Met	Ile	Arg	Thr	Arg	Thr	Ile	Gln	Met	Glu	Pro	Gln
			20					25					30		
Phe	Gly	Gly	Ala	Pro	Cys	Pro	Glu	Thr	Val	Gln	Arg	Lys	Lys	Cys	Arg
			35				40					45			
Ala	Arg	Lys	Cys												
			50												

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 53 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Cys	Arg	Met	Arg	Pro	Trp	Thr	Ala	Trp	Ser	Glu	Cys	Thr	Lys	Leu	Cys
1				5					10					15	
Gly	Gly	Gly	Ile	Gln	Glu	Arg	Tyr	Met	Thr	Val	Lys	Lys	Arg	Phe	Lys
			20					25					30		
Ser	Ser	Gln	Phe	Thr	Ser	Cys	Lys	Asp	Lys	Lys	Glu	Ile	Arg	Ala	Cys
		35					40				45				
Asn	Val	His	Pro	Cys											
			50												

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

[illegible]

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 506 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GAATTCGGCA	NAGGNNAAAC	CCCAGCCCGG	CTGCCGCCCT	GGGCAAGGCC	TNCTGCGCTC	60
TCCTCCTGGC	CACTCTCGGC	GCCGGCACCA	GCCTCTTGGG	GGAGAGTCCA	TCTNTTCCGC	120
CAGAGCCCCG	GCCAAATACA	GCATCACCTT	CACGGGCAAG	TGGAGCCAGA	CGGCCTTCCC	180
CAAGCAGTAC	CCCCTGTTCC	GGCCCCCTGC	GCATGGTNTT	CGCTGCTGGG	GGCCGCGCAT	240
AGCTCCGACT	ACAGCATGTG	GAGGAAGAAC	CAGTACGTCA	TAAACGGGCT	GCGCGACTTT	300
NCGGAGCGGC	GAGGCCTNNG	NCGTTGATGA	AGGAGATCCG	GGNGGCGGGG	GAGGCGTNCA	360
ANAGGTGNCA	AGAGTTNTTT	TCGGGGCCCG	GTTCCCCAAN	GGNAACNGGN	AAACGTTGGG	420
GGNTTTNNAG	TTTNAAGAAG	NAATTNTTGG	TTTTTTTTTTG	GGTGGGATTT	TNCCAACCCN	480
ATTGTTTNTG	GGNTGGAAAA	TTNGAC				506